

TARTU UNIVERSITET
Humanistiska och kontsnärliga fakulteten
Institutionen för främmande språk och kulturer
Avdelningen för skandinavistik
Svenska språket och litteraturen

Kätlin Kulbin

**Guzz, sued, habibi - en kvantitativ analys av ord från invandrarspråk i
svensk skriftspråk**

Bacheloruppsats

Handledare: Maiu Elken

Tartu 2018

Innehållsförteckning

1. Inledning	3
1.1 Invandrasvenska	3
1.2 Invandrarsvenska påverkan på svenska språk	4
1.3 Syfte, begränsning och frågeställning	6
2. Tidigare forskning.....	7
2.1 Teori.....	8
2.1.1 Korpuslingvistik	8
3. Metod och material	10
3.1 Metod.....	10
3.1.1 Språkbankens korpusverktyg	10
3.1.2 Korpusar som har använts för att genomföra undersökningen.....	11
3.2 Material.....	12
4. Analys	15
4.1 Punkter som kommer att besvaras vid analys av varje ord.....	16
4.2 Undersökning av utvalda ord.....	16
4.3 Resultat	24
4.4 Diskussion	26
Sammanfattning	29
Referenser.....	30
Bilagor	32
Bilaga 1	32
Bilaga 2.....	33
Resümee	35

1. Inledning

Språket är ett stort del av varje människans liv. Genom språket kan man uttrycka sig och även måla en bild av sig med hjälp av ord. Vårt ordval kan visa vilken kultur eller vilket socialsamhälle vi hör till. Liksom vårt samhälle ändras, har språket ändrats och det kommer att ändras ständigt genom tiderna. På grund av globalisering och inflytande från media och andra kulturer kan ett språk ändras även snabbare nuförtiden än det har gjort det förut.

Språksituationen i Sverige är ett intressant forskningsämne därför att det redan finns mycket stora mängder av människor i Sverige som har ett annat modersmål än svenska. Förutom finska har arabiskan blivit ett väldigt stort språk i Sverige vilket kan leda till att svenskan blir mer påverkad av arabiskan i framtiden. Samt finns det alltid engelskan som redan har ett stort inflytande på alla andra språk genom sin status som "lingua franca". Eftersom Sveriges population redan är ganska mångkulturell kan man ana att svenskarna kanske är mer öppna till att blanda språket de använder med främmande drag.

1.1 Invandrasvenska

Rinkebysvenska, Albysvenska, Gårdstenska, invandrarsvenska, multietniskt ungdomsspråk och blattesvenska – alla betyder samma företeelse i det svenska språket. De betecknar en blandning av flera språk som invandrare använder för att ibland visa sin tillhörighet eller nämna något de inte kan. En av de största anledningarna till denna företeelse är en ökande invandring till Sverige. *"Ordet "rinkebysvenska" började dyka upp i massmedia på 1980-talet bland annat i samband med språkforskaren Ulla-Britt Kotsinas arbeten".* (Frarud 2003: 82)

Kari Frarud som är Professor emeritus på Stockholms Universitet vid Institutionen för svenska och flerspråkighet hävdar i sin tidskrift att användning av ett annat språk heller är ett val man gör än något som kommer från talarens ursprung.

"Ur ett inifrånperspektiv är alltså rinkebysvenska inte en varietet som talas av alla som bor i Rinkeby, inte ens av alla ungdomar där, och inte heller av alla med invandrarbakgrund eller av andraspråksinlärare eller personer med brytning. Istället är rinkebysvenska ett gruppsspråk som - i varierande grad - talas av ungdomar med eller utan invandrarbakgrund

som bor i flerspråkiga bostadsområden eller är nära vänner till dem som bor där". (Frarud 2003: 84-85)

Samtidigt påstår Ulla-Britt Kotsinas i sin bok "Invandrarsvenska" att de ord som kommer från förortssvenska områden kan sprida sig till andra områden. Till exempel via unga rappare som är ganska uppskattade över hela landet och inte bara i förorten. Hon fortsätter att det händer att även barn som bor utanför förortsområdet leker att de pratar Rinkebysvenska och att "det nya språket" verkar vara intressant för ungdomar och barn som inte bor i förorten där ordet verkligen används. (Kotsinas 2005: 247)

Det samma företeelse där man växlar mellan flera språk kan betecknas med termen kodväxling som används inom språkvetenskapen. Det betecknar fenomenet att språkbrukare ibland byter språk eller dialekt mitt i ett samtal eller i en mening. Vanligtvis består det av utbyte av enstaka ord, oftast substantiv. Kodväxlingen kan användas i flera olika tillfällen och på något sätt kan växlingen mellan olika språk ge en extra dimension och antydan hur språket ska förstås i sammanhanget. Enligt flera forskare kan användningen av flera språk samtidigt räknas som ett språk i sig. (Lindberg 2003: 99-100)

1.2 Invandrarsvenska påverkan på svenska språk

Ulla-Britt Kotsinas påstår i sin bok det att det inte är någonting unikt för Sverige att det förekommer nya varieteter i ungdomsspråket bland ungdomar som har invandrarbakgrund. Hon lyfter fram en undersökning som har gjorts i Paris invandrarrika områden och säger att situationen med ungdomsspråket är likadan i Paris som det är i Sverige. *"Där använder ungdomarna liksom i Sverige ord från sina hemspråk, ofta olika arabiska dialekter, som gemensamma slangord."* (Kotsinas 2005: 247). Hon förklarar att sådan blandning av ord fungerar som ett "hemligt språk" i Frankrike, men liksom i Sverige, sker det en spridning av sådana ord från förorten till mera enspråkiga medelklassområden. Kotsinas fortsätter och säger att det har blivit ganska populärt att använda slangord från förorten i Frankrike och därför försöker konstnärer och författare ta reda på sådana slangord och även andra drag som är kännetecknande för förortsspråk. (Kotsinas 2005: 248)

Kotsinas har förklarat i sin bok att man sedan längre har tänkt att en språkförändring sker av två olika skäl. Först på grund av „inre“ mekanismer i ett språk som analogibildningar och uttalslätnader eller kan ett språk ändras genom lån från andra dialekter eller språk. Men

genom senare undersökningen har det visat sig att processen för språkändringen är mer koplicerad än det var tänkt förut. Bland annat har det visat sig att språkförändringen kan sättas igång med påverkan av utomspråkliga faktorer. Såsom demografiska och sociala förändringar i samhället. Samtidigt har det visat sig att olika sociala grupper spelar en viktig roll i denna språkförändringsprocess. Det har visat sig att förändringarna har spritt sig från grupper med hög social status till grupper med lägre social status. Till exempel de surrande i- och y-ljuden från förorterna med högre status till andra förorten med lägre social status. Men det har hänt även att språkliga drag har spritt sig från en förort med låg social status till en förort med högre social status. (Kotsinas 2007: 167-168) Hur hamnar de nya orden och de språkliga dragen till ett språk? Ulla-Britt Kotsinas påpekar att ungdomar är förmedlare av olika språkliga drag. Med deras hjälp sprids det slangord från invandrargrupperna till homogena svenska grupper. Hon förklarar att det händer eftersom ungdomar har kontakter både i invandrargrupperna och de svenska grupperna. (Kotsinas 2007: 177) Kommunikationen mellan olika grupper med olika bakgrunder är det som hjälper till att nya språkliga drag får en större publik. Nuförtiden krävs det inte heller att fysiskt vara på plats för att prata med andra. Mestadels sker kommunikationen långt från varandra via olika social media plattformar.

Även om Ulla-Britt Kotsinas påstår att det idag finns mer personer som är födda utanför Sverige och det finns mer språk som talas i Sverige än det var under en tidigare period säger hon att det inte betyder att svenskan förstörs eller kommer att försvinna på grund av invandringen. Men hon erkänner att det är sannolikt att en del av slangord som talas i förorterna kan sprida sig utifrån dessa områden. (Kotsinas 2005: 252) Samtidigt påstår Ulla-Britt Kotsinas att svenskan kommer att förändras i framtiden men det betyder inte att alla svenskar kommer att prata "som invandrare". *"Det är visserligen sannolikt att svenskan går inte helt opåvarekad ur en period där en åttondel av invånarna har invandrarbakgrund. Men det forefaller troligt att de förändringar som har sitt ursprung i invandrarungdomarnas varieteter i längden blir ganska marginella och främst beror sådana grammatiska delsystem som redan är i gungning, samt att vissa lexikala element, som idag betraktas som slang, blir integrerade i språket"*. (Kotsinas 2007: 178) Hon erkänner att det är svårt att förutse språkets utveckling eftersom det samtidigt beror på demografiska, sociala och andra utomspråkliga faktorer. Även om vi hör många nya drag hos ungdomar för första gången är det inte säkert att dessa nya drag kommer att stannas kvar i språket. (Kotsinas 2007: 178)

1.3 Syfte, begränsning och frågeställning

Syftet med denna uppsats är att undersöka förekomst av utvalda ord i korpustexter. Ord som jag kommer att undersöka kommer ursprungligen från arabiska, turkiska, syrianska och persiska. I min uppsats vill jag inte fokusera på engelskans påverkan utan ta reda på mer om ord från andra språk har etablerat sig i svenska språket.

Ämnet är relevant från språkforskningperspektiv därför att i vanliga fall utförs undersökningar av likadana ämnen bara på talspråk. Samtidigt finns det undersökningar där det har tagits en skriftlig källa som grunden för undersökningen. Till exempel en roman där författaren har använt ord och uttryck från främmande språk med flit. Oftast för att skapa en bild av den romanfiguren som ligger i centrum eller för att måla en bild av stämningen med hjälp av sådana uttryck. Sedan har undersökningen utförts bara på den källan. Men enligt min mening visar den inte den verkliga situationen som finns i språkanvändningen.

Däremot i den här uppsatsen är antalet förekomster av varje ord det som hela undersökningen baserar på. Jag kommer att fokusera på utvalda ord och undersöka hur ofta och i vilket sammanhang de mest förekommer i utvalda korpustexter. Samt vill jag diskutera deras status i skriftspråket i senare delen av uppsatsen.

Frågeställningen blir således:

- Vilka utvalda ord är mest förekommande i utvalda korpustexter?
- I vilket sammanhang används de mest?

Den här uppsatsen är indelad i följande fyra större huvuddelar. I den första delen ges inledningen till forskningsämnet samt syftet med undersökningen och frågorna som skall besvaras i uppsatsen. I den andra delen redovisas tidigare forskning och förklaras teorin som ligger bakom undersökningen. I den tredje delen berättas om metoden och materialet analysen baserar på. I den fjärde delen analyseras undersökningsmaterialet. Resultat kan ge mer information om språkets utveckling samt visa om ord från andra språk kan sprida sig i det svenska skriftspråket.

2. Tidigare forskning

Detta avsnitt kommer att ge en kort översikt av tidigare forskning gjort av samma ämne. Språkets utveckling har varit ett intressant ämne för flera forskare. En av de mest kända språkforskarna som har ägnat sig åt att undersöka slang och invandrarsvenska är Ulla Britt Kotsinas. Hon var också medlem i en bred undersökning av slangord genomförd i alla Nordiska länderna med syftet att studera språkkontakters inflytande på alla Nordiska språk.

En stor undersökning var genomförd under åren 1996 - 2000 av projektet Språkkontakt och Ungdomsspråk i Norden (UNO). Syftet med denna undersökning var att studera ungdomsspråket i de Nordiska länderna och ta reda på hur språket påverkas genom språkkontaktarna. Resultaten är presenterade i en rapport "Ungdommers språkmöter" av A.B Stenström, U.B Kotsinas och E.M Drange. Även om undersökningen handlar om alla Nordiska länder kommer jag att ge bara en översikt av den Svenska delen av undersökningen. För att genomföra undersökningen av slangord i Sverige ombads ungdomar mellan 15-17 års ålder i en skriftlig enkät att omge så många slangord de kunde för sammanlagt 55 normalord. I den Svenska delen av undersökningen deltog arton gymnasieskolor och nio grundskolor och sammanlagt besvarades enkäterna av 2 105 ungdomar i Sverige. I resultaten visades det att förutom engelskan kommer de flesta slangorden och lånorden från turkiska, arabiska, assyriska/syriska och spanska. Respektive 203, 62, 49 och 82 lexikonord för varje språk.

Språk	Antal lexikonord	Antal belägg
turkiska	203	1294
arabiska	62	278
assyriska/syrianska	49	186
spanska	84	180
grekiska	28	48
italienska	11	33
kurdiska	15	25

(Stenström, Kotsinas, Drange 2000: 39-40)

2.1 Teori

Detta avsnitt kommer att förklara vilka metoder som används för att besvara undersökningsfrågorna. Uppsatsen ska huvudsakligen undersöka förekomst av ord som ursprungligen kommer från arabiska, turkiska och assyriska/syrianska i utvalda korpustexter.

2.1.1 Korpuslingvistik

Den största delen av analysen i den här uppsatsen sker via metoden som kallas för korpuslingvistik. Det är en metod där forskaren har använt ett korpusverktyg för att genomföra undersökningen och samla data. Tony Mcenery och Andrew Hardie definierar korpuslingvistiken på följande sätt. Vi kunde lätt definiera korpuslingvistik som en sätt att hantera en samling av maskinläsbara texter som anses vara en lämplig grund för att studera en specifik uppsättning av forskningsfrågor. (McEnery 2012: 1) *“I och med att man i större skala började använda sig av datakorpusar etablerades termen computer corpus linguistics eller datakorpuslingvistik, vanligtvis kallad korpuslingvistik, som beteckning för metoden.”* (Rawoens 2003:2)

En korpus är en samling av data och enligt Graeme Kennedy som är professor i tillämpad forskning inom lingvistik kan en elektronisk korpus bestå av kontinuerliga exempel av texter, vilka tas fram från hela texter eller kan även bestå av samlingar av citat. Han fortsätter och förklarar att korpusar har utnyttjats för flera olika skäl. Till exempel är några nuvarande korpusar avsedda för språklig forskning och de har utformats så att de kan undersökas för att svara på frågor på olika språkliga nivåer såsom prosodier, grammatik, lexikala elementer, pragmatik eller diskursmönster.

Andra korpusar kan vara utformade för olika specialiserade ändamål. Först kan de användas för att upptäcka vilka ord och ordbetydelser som borde ingå i en lärares ordbok. Eller kan korpusar användas för att ta reda på vilka ord eller ordbetydelser mest används av arbetstagaren inom ekonomin eller oljeindustrin och även vilka skillnader det finns mellan användningen av språk i olika sociala, geografiska, arbetsrelaterade eller historiska sammanhang. (Kennedy 2012: 3-4)

Vikten av korpusar är beskriven igen av samma forskare på följande sätt:

En korpus utgör den empiriska grunden, inte bara för att identifiera de element och strukturmönster som utgör de system vi använder på ett språk, men också för att kartlägga vår användning av dessa system. (Kennedy 2014: 3) Gudrun Rawoens har förklarat den autentiska grunden till denna metod i sin artikel: *“I språkforskning som bedrivs inom korpuslingvistik utgör en korpus alltså den empiriska basen”*. (Rawoens 2003: 1) Men samtidigt är det viktigt att antingen det utgörs en kvalitativ eller en kvantitativ undersökning beror vikten av analysresultat på faktorn att de korpusdata vi väljer för att utforska en frågeställning, måste passa med den frågan som ställs i början. (McEnery 2012: 2) Graeme Kennedy fortsätter att en korpus kan analyseras och jämföras med andra korpusar eller andra delar av korpusar för att studera variation. Viktigast är att en korpus kan analyseras distribution av ord för att visa hur ofta särskilda fonologiska, lexikaliska, grammatiska, diskursala eller pragmatiska särdrag uppträder, och även där de förekommer. (Kennedy 2014: 3)

Enligt Tony McEnery och Andrew Hardie är korpuslingvistik inte en gemensam, monolitisk överenskommen uppsättning av metoder och förfaranden för utforskning av ett språk. Medan vissa generaliseringar kan göras som karakteriserar mycket av det som kallas "corpus linguistics" är det väldigt viktigt att inse att korpuslingvistik är ett heterogent fält. Skillnader finns inom korpuslingvistik som skiljer ut och subkategoriserar olika tillvägagångssätt för användningen av korpusdata. (McEnery 2012: 1) Användning av korpus för att samla undersökningsdata är en föredragen metod hos olika forskare som redan Charles F. Meyer påstår i sin bok. *“Även om beskrivande / teoretiska lingvister och datorlärdare använder corpora för mycket olika ändamål, delar de en gemensam övertygelse: att det är viktigt att basera en analys av språk på verkliga data - faktiska instanser av skriftligt skrivande - i stället för data som är förbryllad eller "uppbyggd"”*. (Meyer 2002: xiii)

3. Metod och material

3.1 Metod

Både kvantitativ och kvalitativ metod används för att besvara undersökningsfrågorna.

Undersökningen är kvantitativ därför att det kommer att hittas svar på följande punkter vid analysen av varje ord. Jag kommer att ta reda på antal sökträffar, förekommande lemgram, sammanhang där ord förekommer mest i och under vilket år ordet först förekom i korpustexter för varje utvalda ord.

Undersökningen är kvalitativ därför att det efteråt jämförs både orden och resultaten. För att hitta gemensamma mönster hos dem. Resultaten uttrycks med hjälp av ord.

3.1.1 Språkbankens korpusverktyg

Språkbanken är en forskningsenhet vid Institutionen för svenska språket på Göteborgs universitet. Deras lexikonresurser är sökbara online och fritt för alla typer av användning. (Språkbanken) På det här sättet gör det lättare för en forskare att samla data. Språkbankens Korp innehåller naturligt språkbruk och samtidigt autentiska texter som helst är representativa för en texttyp, genre, författare eller tidsperiod. Genom att använda en korpus kan forskaren se hur språket faktiskt används.

Undersökningen är baserad på totalt 167 korpusar av 237. 13 av korpusarna är skyddade och andra korpusar valde jag inte att kryssa i för att få en relevant statistik över materialet. Medan jag genomförde min undersökning bestod alla utvalda korpusar av 11,25G token. Under fliken statistik har jag valt att sammanställa sökträffarna på ord, ordklass och lemgram. Ett lemgram innefattar alla böjningsformer som ett ord kan ha. Om det till exempel söks efter ordet *katt* visas det även former som *katter* och *katterna* bland alla träffar. På det här sättet är det även möjligt att se vilka de mest förekommande böjningsformerna är, ifall det finns några. Med alla möjliga ordklasser visas det om ordet kan användas i en annan funktion. I fall andra möjliga betydelser förekommer, räknas det med i analysen bara det antalet hur många gånger ord i samma ordklass som jag har markerat i tabellen där alla ord som kommer att analyseras finns med används (se tabell 1).

För att känneteckna ord, som förekom i andra möjliga betydelser enligt Korp, har jag noterat det i analysdelen vid varje ord. Om jag har räknat bara antal förekomst av ett ordklass då finns det med i spalten "antal träff" punkt. I fall ordet har inte förekommit i andra betydelser, finns det ingen noteringen i analysen heller. Vid undersökningen om hur ofta ord har uppstått årligen har jag använt valet „visa trenddiagram“ under fliken „statistik“:

3.1.2 Korpusar som har använts för att genomföra undersökningen

Meningen är att förklara från vilka källor texter har samlats ihop.

Akademiska texter (humaniora, samhällsvetenskap)

De består av två korpusar med akademiska texter.

Finlandssvenska texter (webbtexter, skön- och facklitteratur, tidningstexter, tidskrifter)

Texter från finlandssvensk språkbank.

Skönlitteratur

Innehåller texter från:

Aspac svenska, Bonniersromaner I (1976-77), Bonniersromaner II (1980-81), Norstedtsromaner (1999), SUC-romaner, äldre svenska romaner.

Sociala medier (bloggmix, diskussionsforum, Twitter)

Blogg mix från år 1998 fram till 2017, samtidig innehåller texter med okänt datum. Här finns texter från ett urval av svenska bloggar. Korpus uppdaterades regelbundet.

Det finns två diskussionsforum från vilka texterna har blivit samlade. En av dem kallas för Familjeliv och den andra Flashback forum. Korpus uppdateras regelbundet.

Twitter mix består av material från ett urval av svenska Twitter användare. Korpus uppdateras regelbundet.

Tidningstexter (GP, Press)

Urval av texter från Göteborgsposten från år 1994 och från 2001 fram till 2013.

Urval av texter från svensk press år 1965, 1976, 1995, 1996, 1997, 1998

Webbnyheter (8 sidor)

Nyhetsartiklar från 8 sidor.

Svensk Wikipedia

Artiklar från svenska Wikipedia. Korpus uppdateras regelbundet.

3.2 Material

I det här kapitlet presenteras det utvalda materialet som används för att genomföra undersökningen. Analysen baseras på ett antal ord från ovannämnda invandrarspråk som är slumpvis samlade från olika källor, till exempel Svensk Slang Ordbok på nätet, Slangopedia och Nyrordslistor som ges ut varje år av Institutet för Språk och Folkminnen. Svensk Slang Ordbok och Slangopedia har samtidigt använts för att ta reda på varifrån varje ord ursprungligen har kommit.

Från språkforsknings perspektiv är det ett relevant forskningsämne på grund av ökandet av antal invandrare i Sverige som räknar sig att ha ett annat modersmål än svenska. Den aktuella språksituationen har varit ett intressant undersökningsämne för flera forskare, en av dem är Mikael Parkvall. Han är språkforskare vid Institutet för lingvistik på Stockholms Universitet. Han har undersökt språksituationen i Sverige och givit ut flera böcker där han har samlat data hur många talare det finns för varje främmande språk i Sverige. Till exempel "Sveriges Språk - Vem talar vad och var", utgiven år 2009. En senare utgåva av samma forskare kallas för "Sveriges språk i siffror: Vilka språk talas och av hur många?" var utgiven år 2016. Här har jag jämfört hans undersöknings resultat.

I tabellen nedanför har jag utvalt några relevanta exempel för Sveriges språksituation år 2009 och 2012 från ovannämnda böcker. Jag har nämt språket och sedan dess användare som räknar språket som sitt modersmål. I dessa böcker hade Mikael Parkvall följande statistik för invandrarspråkens användning i Sverige.

Språk	Antal talare (2009)	Antal talare (2012)
Serbokroatiska (serbiska, kroatiska, bosniska)	113 000	155 000
Arabiska	93 000	130 000
Kurdiska	~ 66 000	84 000
Persiska	59 000	74 000
Turkiska	34 000	45 000

(Parkvall 2009: 90-95)

I den senare undersökningen gjord av Mikael Parkvall kan man se att antal användare har ökat. I den senare undersökningen som var utgiven år 2016 gäller siffrorna för år 2012. Jag kunde ta fram materialet genom en annan källa och lägga det fram här i undersökningen för en komparativ syfte.

Det visas från Parkvalls undersökning att antal talare för varje invandrarspråk ökas ständigt. Det finns redan väldigt många i Sverige som räknas att ha ett annat modersmål än svenska. På grund av det är det förväntat att det finns ganska många ord som har ett främmande ursprung i folkets ordförråd. För min egen undersökning har jag valt ord från olika källor på internet. Alla ord är utvalda baserade på mitt eget intresse beroende på vilka av de har verkat vara mest intressant att ta reda på mer.

I tabellen nedanför har jag nämt alla ord som analysen baserar på för att besvara undersökningsfrågorna. Jag har urskiljt ord från varandra på grund av deras ordklass, ursprung, betydelse och från vilken källa de har tagits ifrån. Alla orden har kategoriserats i alfabetisk ordning.

ORD	ORDKLASS	URPSRUNG	BETYDELSE	KÄLLA
abri	substantiv	syrianska	bror / kompis	Slangopedia
albi	substantiv	arabiska	älskling	Slang Ordbok
araba	substantiv	turkiska	bil	Slangopedia
balesh	adverb	turkiska, kurdiska	gratis	Slang Ordbok
benim	substantiv	turkiska	jag själv	Slang Ordbok
brakkam	substantiv	kurdiska	bror	Slangopedia
dunya	substantiv	arabiska	värld	
flos	substantiv	arabiska	pengar	P5 STHLM
guzz	substantiv	turkiska	tjej	Slang Ordbok
gäri	substantiv	turkiska	tjej	Nyordslista 2014
habibi	substantiv	arabiska	älskling	Slang Ordbok
haffla	verb/substantiv	arabiska	festa/fest	GP
jalla	verb	arabiska	skynda	Slang Ordbok
parra	substantiv	turkiska/ kurdiska	pengar	Slang Ordbok
salaam	substantiv	arabiska	fred	P5 STHLM
suedi	substantiv	persiska	svensk person	GP / Slang Ordbok
shaytan	substantiv	arabiska	djävul	Slangopedia
shuno	substantiv	tyrkisk, assyrisk, syriansk	kille	Svensk/ dansk ordbok
tamam	adjektiv	syrianska	okej / Bra	Slangopedia
walla	verb	arabiska	svär på Gud	Slangopedia

Tabell 1. Ord som kommer att analyseras

4. Analys

I detta kapittel analyseras alla ord som är samlade ihop för att genomföra undersökningen. Jag har gjort sökningar på alla ovannämnda ord med hjälp av Språkbankens korpusverktyg Korp, för att se hur ofta varje ord förekommer i utvalda korpustexter samt i vilket sammanhang de mest används, vilka lemmar som finns för varje ord och vilket år de förekom i texten för första gången. Metoden att använda korpusar för språkforskning kallas för korpuslingvistik.

Genom analysen vill jag svara på följande frågor:

- Vilka utvalda ord är mest förekommande i de utvalda korpustexterna?
- I vilket sammanhang används de mest?

För att genomföra undersökningen har jag använt Korp version 7.0.0. Jag har använt en enkel sökning i korpusverktyget. Undersökningen påbörjades den 10.05 och varade fram till 16.05. Följande korpustexter används för att genomföra undersökningen:

- Akademiska texter (humaniora, samhällsvetenskap)
- Finlandsvenska texter (webbtexter, skön- och facklitteratur, tidningstexter, Tidskrifter)
- Skönlitteratur
- Sociala medier (bloggmix, diskussionsforum, twitter)
- Tidningstexter (GP, Press)
- Webbnyheter (8 sidor)
- Svensk Wikipedia

Vid analysen av varje ord har jag lagt till svar på följande frågor. Samt gav några autentiska exempel på deras förekomst för att illustrera den möjliga kontexten orden kan förekomma i texterna. En tabell, där antal förekomst av varje ord i alla utvalda korpustexterna finns med, hittas i bilagor. Se bilaga 1. En tabell som illustrerar förekomst av ord per år, hittades också i bilagor. Se bilaga 2.

4.1 Punkter som kommer att besvaras vid analys av varje ord

- **Antal träffar**

Antal träffar har hittats från Språkbankens korpusverktyg. Det står under fliken KWIC och den står för "keyword I context".

- **Förekommande lemgram av varje ord**

Ett lemgram innefattar alla böjningsformer som ett ord kan ha. På det här sättet är det möjligt att se vilka är den mest förekommande böjningsformen av varje ord.

- **Sammanhang**

Det betyder att det har noterats i vilken text ordet mest förekommer. Genom detta kan man se var ordet mest används.

- **Första belägg**

Det betecknas för år varje ord när det har förekommit i korpustexterna för första gången

4.2 Undersökning av utvalda ord

Abri (bror, kompis)

Antal träffar: Totalt förekom ordet 71 gånger i de utvalda korpustexterna.

Lemgram: Inga träffar av andra böjningsformer.

Sammanhang: Förekom mest diskussionsforum texter, totalt 34 gånger.

Första belägg: Förekom för första gången år 2002.

Exempel:

"Grattis på födelsedagen abri!" (Twitter)

"Välkommen Özcan, nu kör vi abri!" (Twitter)

"Vad snackar du om abri?" (Flashback: Samhälle)

Albi (älskning)

Antal träffar: Ordet förekom som substantiv totalt 85 gånger i de utvalda korpustexterna.

Lemgram: Gav inga träffar.

Sammanhang: Ordet förekom mest i diskussionsforum texter, totalt 42 gånger.

Första belägg: Förekom för första gången i en bloggtext år 2005.

Exempel:

”Jag kallade honom för albi för SD och att jag inte avskyr SD:are, men jag avskyr de SD:are med invandrarbakgrund som är rasister och går emot sina ”egna”. (Blogg 2011)

”nej men tack albi <3” (Twitter)

“albi krya på dig, hade velat se ditt anz ikväll” (Twitter)

Araba (bil)

Antal träffar: Ordet förekom som substantiv totalt i 85 gånger.

Lemgram: Gav inga träffar.

Sammanhang: Förekom mest i Twitter, totalt 78 gånger.

Första belägg: Ordet förekom för första gången år 2005.

Exempel:

”Asså jag e så fkn borta på mornarna, satt i bussen i 20 mins tills jag ba ey vart är mitt bälte trodde jag va i en araba bre.” (Twitter)

”Trodde jag skulle få gå ut nu när jag har araba men nej.” (Twitter)

Balesh (gratis)

Antal träffar: Ordet förekom totalt 2 gånger

Lemgram: Gav inga träffar

Sammanhang: Förekom en gång i Familjeliv: Allmänna rubriker – Samhälle och en gång i Twitter.

Första belägg: Ordet förekom för första gången i diskussionsforum år 2014.

Exempel:

”Tackar inte nej till saker som e balesh” (Twitter)

Benim (jag själv)

Antal träffar: Ordet förekom som substantiv totalt 265 gånger.

Lemgram: Gav inga träffar.

Sammanhang: Förekom mest i Twitter, totalt 254 gånger.

Första belägg: Ord förekom för första gången år 2005.

Exempel:

"1månad benim är fri, som en "fågel" #ingenstress, #ingenstress" (Twitter)

"Men benim bryr sig den här gången bre" (Twitter)

Brakkam (bror)

Antal träffar: Förekom totalt bara 2 gånger

Lemgram: Gav inga träffar.

Sammanhang: Förekom bara i Twitter, totalt 2 gånger.

Första belägg: Ord förekom för första gång år 2017.

Exempel:

"Välkommen brakkam" (Twitter)

Dunya (värld)

Antal träffar: Substantivet förekom totalt 63 gånger i de utvalda korpustexterna.

Lemgram: Gav inga träffar.

Sammanhang: Förekom mest i Twitter, totalt 41 gånger.

Första belägg: Förekom för första gången år 2007.

Exempel:

“Sedan inser de att det också är en del av dunya och istället försöker lyssna till Skaparen av denna dunya och finna njutning och ro från skapelsen genom att fly till Skaparen.” (Blogg 2009)

“Islam för mig har hjälpt mig att finna sanningen och kärleken på denna “dunya”. (Flashback: Vetenskap & Humaniora)

Flos (pengar)

Undersökningen med det här ordet var mer problematisk därför att ordet *flos* kan även förekomma i en annan betydelse. Till exempel finns det en inlägg i Göteborgs Posten 1994 – “Florence har bildat av latinets *flos*, som betyder blomma.”

Antal träffar: Förekom totalt 153 gånger i korpustexter, i en kontext där det betyder pengar.

Lemgram: Inga relaterade träffar.

Sammanhang: Ordet förekom mest i diskussionsforum, totalt 88 gånger.

Första belägg: Ordet förekom för första gången i diskussionsforum år 2001.

Exempel:

“Hoppas verkligen att det inte är något annat som är trasig som kostar flos.” (Flashback: Fordon & Trafik)

"Alla vet inga flos i musiken." (Twitter 2012)

"Badar i flos men bor ändå fult." (Twitter 2012)

Guzz (tjej)

Resultat med detta ord är annorlunda därför att det var det enda ordet för vilket jag kunde hitta lemmgram. Eftersom alla resultat var markerade som substantiv i korpus texter räknade jag ihop alla träffar.

Antal träffar: Totalt förekom ordet i texterna 2 556 gånger.

Lemgram: Vanligaste lemmgram som förekom i texterna var ordets grundform, *guzz*, vilket förekom 1 618 gånger totalt.

Andra förekommande lemmgram: *guzzar* – förekom 498 gånger, *guzzen* – förekom 257 gånger.

Sammanhang: Ordet förekom mest i Twitter – totalt 1 599 gånger.

Första belägg: Ordet förekom för första gången år 1997 i en GP utgåva.

Exempel:

“Ibland skriver han dikter till henne, den vackraste guzzen.” (Bloggmix 2007)

”Ok en fett nice dude klädd som guzz.”

”Vi kan göra språket rikare genom att ta in nya ord som keff och guzz.”

Gäri (tjej)

Antal träffar: Ordet förekom som substantiv i texterna totalt 1 665 gånger.

Lemgram: Gav inga träffar.

Sammanhang: Ordet förekom mest i Twitter, totalt 1 625 gånger.

Första belägg: Ordet förkom för första gången i diskussionsforum år 2011.

Exempel:

”Ska bli så sjukt kul att få hänga med min gäri gen!” (Bloggmix 2013)

"Din gäri är värd ett telefonsamtal" (Twittermix)

"Tvätta bilen hämta gäri, äta middag o sen ut o gå" (Twitter 2015)

Habibi (älskling)

Antal träffar: Totalt förekom ordet 2 201 gånger i de utvalda korpustexterna.

Lemgram: Gav inga träffar.

Sammanhang: Ordet förekom mest i Twitter, totalt 1 468 gånger

Första belägg: Första belägg hittades från SVD från år 1997.

Exempel:

"Heej habibi <333 vad gör du?" (Blogg 2005)

"Ska snart iväg och hämta min habibi från dagis, tiden rinner iväg, han fyller två nästa vecka, hamdullah, han är värt allt." (Familjeliv - Allmänna rubriker - Samhälle 2006)

Haffla (festa, fest)

Antal träffar: Ordet förekom som verb och substantiv i de utvalda korpustexterna totalt 290 gånger.

Lemgram: Sökningen gav inga träffar.

Sammanhang: Förekom mest i Twitter, totalt 277 gånger.

Första belägg: Första belägg hittades från år 2007 i en blogg.

Exempel:

"Skivan blandar hiphop, dancehall, och pop till väldigt festliga toner vars blandning kallas för "haffla music". (GP 2013)

"Ingen annan kommer ju fatta när jag säger "toch toch toch eyy mannen ska vi haffla i helgen??" (Blogg 2007)

"#haffla ikväll" (Twitter)

Jalla (skynda)

Antal träffar: Totalt förekom ordet som verb i de utvalda korpustexterna 1 461 gånger.

Lemgram: Gav inga träffar.

Sammanhang: Ordet förekom mest i Twittermix, totalt 716 gånger

Första belägg: Först inträffades år 2000.

Exempel:

"Jalla jalla, en fet kebab till alla" (blogg 2011)

"Bara två kvar.. jalla jalla skynda er nu ni sista två så vi kommer igång *nyfiken och otålig*" (Familjeliv: Allmänna rubriker)

Parra (pengar)

Antal träffar: Ordet förekom som substantiv totalt 124 gånger i de utvalda korpustexterna.

Lemgram: Gav inga träffar.

Sammanhang: Förekom mest i diskussionsforum texterna, totalt 57 gånger.

Första belägg: Först förekom år 2003

Exemplar:

“Lizette: Jag har lunch snart, men jag har knappt på parra så jag äter nog hemma. Linnea: Så trist, jag hade velat fara på Simonssons. Lizette: Åh, okej då!” (Blogg 2009)

“Alltså önskar att parra finns, vill tatuera mig igen sooo badly :(“ (Twitter)

Salaam (fred)

Antal träffar: Ordet som substantiv förekom totalt 183 gånger i de utvalda korpustexterna.

Lemgram: Gav inga träffar.

Sammanhang: Förekom mest i diskussionsforums sammanhang, totalt 132 gånger.

Första belägg: Först inträffades i 2005.

Exempel:

„Fred med dig, salaam“ (Flashback: Vetenskap & Humaniora)

Shaytan (djävul)

Antal träffar: Ordet förekom totalt 186 gånger i de utvalda korpustexterna.

Lemgram: Gav inga träffar.

Sammanhang: Förekom i diskussionsforums texter, totalt 105 gånger.

Första belägg: Förekom för första gången i Familjeliv: Övrigt forum i år 2002.

Exempel:

”Slutligen Allt som är rätt i den här artikeln är en gåva från allah och alla misstag är från mig själv eller shaytan!” (Familjeliv forum 2006)

”Om det finns något som är fel beror det på mig och shaytan och det som är rätt är ifrån Allah (tw)” (Familjeliv 2002)

”Asså är det nån slags mode nu att ha shaytan ögonbrun lr?” (Twitter)

Shuno (kille)

Antal träffar: Ordet förekom totalt 1 333 gånger i de utvalda korpustexterna.

Lemgram: Gav inga träffar.

Sammanhang: Förekom mest i Twitter, totalt 1 254 gånger.

Första belägg: Först inträffades år 1997.

Exempel:

”I går när en random shuno kom fram och kopplade in sin telefon i mitt dj-bås?” (Twitter)

” hahahah Malmös tränare var ju en skön shuno” (Twitter)

Suedi (svensk person)

Antal träffar: Totalt förekom ordet i de utvalda korpustexterna 976 gånger.

Lemgram: Bara ordets grundform förekom, sökning med lemgram gav inga träffar.

Sammanhang: Förekom mest i Twittermix. Totalt 884 gånger.

Första belägg: Först inträffades år 2005.

Exempel:

”En historia på ett språk där coola förortsskillar är shunnar och trista svenskar är suedi och där de flesta orden är bortskalade till ett hårt, effektivt staccatospråk.” (GP 2009)

”Idag åt jag potatis och lingonsylt, är jag #suedi nu?” (Twitter)

"Volvo v79 ala suedi" (Flashback: Fordon & Trafik)

Tamam (bra, okej)

Antal träffar: Ordet förekom som adjektiv totalt 172 gånger i texterna.

Lemgram: Gav inga träffar.

Sammanhang: Förekom mest i Twittersammanhang, totalt 157 gånger

Första belägg: Ordet förekom i bloggmix år 2004.

Exempel:

"Annars har jag bara jobbat som vanligt och småfixat med lägenheten en maaaassa FIFA har man oxå hunnit med, många vill ha bilder men det får komma sen då allt är tamam!" (Bloggmix 2004)

"Med lite vanilla flavour blir den tamam."

"Jag hade en tamam dag med er tjejer. Shout out to; Naima, Amal, Aisha, Fatima, Rahima och Maryam S." (Blogg 2012)

“Minns att yoghurtsåsen va tamam” (Flashback)

Walla (svär på Gud)

Antal träffar: Ordet förekom som verb totalt 1 360 gånger i texterna.

Lemgram: Gav inga träffar.

Sammanhang: Verbet förekom mest i Twitter, totalt 1 016 gånger.

Första belägg: Förekom för första gången år 2003.

Exempel:

“Sen walla det finns inte så mycket att tillägga.” (Blogg 2007)

“Ok, då jag kan förklara: Ironi är en sak, sarcasm är en annan jag scar walla, det är så.”

(Flashback: Hem, Bostad & Familj”

4.3 Resultat

I tabellen nedanför finns alla orden med deras betydelse och förekomst i de utvalda korpustexterna.

ORD	ANTAL TRÄFFAR
guzz (tjej)	2 556
habibi (älskling)	2 201
gäri (tjej)	1 665
jalla (skynda)	1 461
walla (svär på Gud)	1 360
shuno (kille)	1 333
suedi (svensk person)	976
haffla (festa, fest)	290
benim (jag)	265
shaytan (djävl)	186
salaam (fred)	183
tamam (okej, bra)	172
flos (pengar)	153
parra (pengar)	124
albi (älskling)	85
araba (bil)	85
abri (bror, kompis)	71
dunya (värld)	63
balesh (gratis)	2
brakkam (bror)	2

Tabell 2. Resultat av analysen

Det visas att ord med det största antal träffar var *guzz* vilket ursprungligen kommer från turkiskan. Näst mest förekommande ord var *habibi* vilket kommer från arabiskan. Tredje mest förekommande ord var *gäri*, som kommer från turkiskan. Ord med det minsta antal träffar var *brakkam*, som kommer från kurdiskan och *balesh*, som kommer från turkiskan. De förekom både endast två gånger i de utvalda texterna.

Med några undantag förekom orden mest i Twitter sammanhang men samtidigt visas det tydligt att orden förekommer ofta även i diskussionsforum „Flashback“ och „Familjeliv“. Sedan med några träffar förekommer ordet i bloggtexter.

Bland första beläggen finns det inget särskilt mönster. Det finns ord såsom *habibi*, *guzz* och *shuno*, som förekom för första gången i texterna väldigt tidigt, redan år 1997. Både *guzz* och *shuno* förekom i GP, där det beskrevs ordens betydelse och ursprung. Däremot „gäri“ är det enda ordet, som förekom ganska sent bland andra. Men dess användning har ökat mest genom åren.

Det enda ordet för vilket det visades andra lemgram var *guzz*. Mest förekommande lemgram var *guzzar* och *guzzen*. De förekom respektive 498 gånger och 257 gånger.

4.4 Diskussion

I det här kapitlet ska det diskuteras över resultaten som jag har fått genom undersökningen. Genom den här analysen kan man få veta hur de orden används i praktiken. Vilka ord används mest i de utvalda texterna och i vilka sammanhang de mest används? Det är förståeligt att främmande ord sprids oftast och samtidigt snabbare via muntligt tal. Men det som känneteknar det att ett ord har blivit etablerat i ett språk är deras förekomst i skriftspråket. Det är antagligt att några lånord endast sprids via talspråket, men det är intressant att få veta om några av dem även har kommit in i skriftspråket. När ett ord blir mer använt i skriften kan det tyda att det har kommit in i själva språket. Även om användningen av ord kan variera. Som diskussion om hela betydelsen av ett främmande ord, ironiseringen, eller när man faktiskt använder ett främmande ord för att beskriva någonting som det redan finns ett ord för i språket.

Det visas en tydlig tendens att nästan alla utvalda orden var mest förekommande i Twitter sammanhang. Det kan antingen tyda på att Twitter används mycket av folket eller att det bara är lättare att använda Twitter för korta meddelanden man vill dela med andra. Samtidigt är Twitter inte så formellt i sin kontext, där är det möjligt att fritt dela sina tankar med alla andra. Vilket gör det möjligt att man använder ord, man vanligtvis inte kunde använda i formella sammanhang. På det sättet är Twitter en ekvivalent till skrivet talspråk. Liksom i muntligt språk sprids främmande ord snabbare i ett fritt, oformellt skriftspråksmedium som Twitter egentligen är i sin grund. Ett annat populärt sammanhang där ord med främmande ursprung förekom tydligt ofta är olika slags rubriker i diskussionsforum. I utvalda korpustexter finns det med två korpusar som innehåller material från forum "Familjeliv" och "Flashback". Ord som *albi*, *abri*, *parra*, *flos*, *salaam* och *shaytan* förekom mest i diskussionsforum. Det kan bero på faktum att olika diskussionsforum har funnits med innan Twitter introducerades. Twitter lanserades år 2006.

Det är antagligen så att det tar tid innan en ny plattform blir populär bland användare. Enligt statistiken har Twitter idag 336 miljoner aktiva användare i jämförelse med år 2010 då det bara fanns 10 miljoner aktiva användare. (Statista) Vad gör det enklare att ord sprids snabbare i Twitter är det att Twitter används över hela världen och man kan dela sina tankar med hela världen. Däremot diskussionsforum såsom "Familjeliv" och "Flashback" är bekanta bara för användare i Sverige. Vilket begränsar deras tillgång till den allmänna publiken.

Samtidigt uppstod det en intressant frekvens. Nämligen ord för "kvinnor" och "man" är mer förekommande i korpustexterna än ord för vanliga saker till exempel "bil". Det kan förklaras med att ord för beteckna människor är mer användbara. De är sådana ord som alla lär sig först när de börjar att tala. Samtidigt kan ord för "man" och "kvinna" användas mer i flera olika kontexter än ett ord för "bil" till exempel. Den här analysen kan inte ge hela bilden av förekomsten av de utvalda orden. Det är därför att de utvalda orden även kan stavas på annorlunda sätt än i den här undersökningen. Till exempel finns det både "Yalla" och "Jalla", "Walla" och "Wallah", "Guss" och "Guzz" som egentligen har samma betydelse men olika stavningssätt. Deras varierande stavningsätt beror på att lånord sprids muntligt i talspråket och det finns inga allmänna regler om hur ett främmande ord borde skrivas. Förekomst av olika stavningsätt beror på hur en användare har valt att stava ett främmande ord. I analysen har jag bara undersökt förekomst av ord med ett stavningssätt.

En annan intressant fynd var att ordet *guzz* förekom med *-ar* och *-en*. Ordet *guzz*, även om det är ett främmande ord, böjs med suffix som är vanliga för svenska språket. Det var alltså det mest förekommande ordet bland mitt urval av analysmaterial och även dess lemmogram förekom ganska ofta de i utvalda korpustexterna. Ordet *guzzar* – förekom 498 gånger, *guzzen* – förekom 257 gånger. Ordets böjningsätt kan tyda på att det kanske är det mest integrerade ordet bland mitt urval av ord från invandarspråken. Ord med samma betydelse men med annorlunda stavningsätt „guss“ har blivit lagt till i Svenska Akademiens Ordlista (SAOL). Bland mina utvalda ord är „*gäri*“ den som har blivit markerat även i Nyordslistan. Ordet var med i Nyordslistan år 2014. Undersökningen visar att „*gäri*“ är det enda ordet som samtidigt uppstod ganska sent i textkorpusarna. Ordet inräffades först år 2011 men efter ordet blev introducerat i Nyordslistan har dess användning ökat kraftigt. Den hade mest träffar år 2017 bland andra utvalda ord. När jag undersökte ord i korpustexterna visas det sig att vid användningen av ord "suedi" träffades ordet ganska ofta i samma kontext. "#Vackna upp och va suedi" Ordet användes ganska ofta på grund av det att en svensk rappare Erik Lundin hade gjort en rap-låt med samma titeln "*Suedi*". Som det visas i analysen förekom ordet ganska ofta i utvalda korpustexterna totalt 976 gånger. Hans EP publicerades år 2015. En annan låt som var med i samma EP har titeln "*Haffla*", som också är med i denna undersökning, men detta ord förekom inte så ofta i en kontext som hänvisade till denna rappare, samt förekom ordet inte så ofta heller som "*suedi*" totalt 290 gånger. Som redan framgått i inledningen påpekar Ulla-Britt Kotsinas att musik kan hjälpa till att sprida främmande ord från

invandrartäta förorter till andra områden. Jag kan inte bevisa att ordet har spritt sig till andra förorter där det inte finns så många invandrare baserade endast på korpustexterna men musiken hörs över hela landet, vilket gör den en effektiv förmedlare av ord.

Vad gjorde det svårare att undersöka ord med främmande ursprung i texterna var det hur ord var kategoriserade i korpustexterna. Även om några ord var kategoriserade så att de kunde förekomma även i en annan ordklass och betydelse, var det svårt att bedömma över det att de faktiskt hade en annan betydelse när man bara utgick från frasen där ordet användes. Till exempel ordet *tamam* som borde betyda "okej eller bra" och vara ett adjektiv förekom enligt Korp även i form av substantiv och det gjorde det mer än i form av adjektiv, respektive 257 och 172 gånger. I meningarna: "Bilderna blir *tamam* med 12Mpix!" och "Skivan igår va *tamam*" räknades ordet *tamam* som ett substantiv. Däremot i meningarna "Vin och broderier är *tamam*" och "Varför är livet så *tamam* just nu?" räknades ordet *tamam* som ett adjektiv. I analysen räknade jag med antal träffar baserade på korpus information och noterade antal träffar där ordet *tamam* förekom i form av adjektiv.

Det är klart att den här uppsatsen inte kan ge en bild av hela situationen om invndrarsvenska drag i svenska språket - eftersom det finns bara 20 ord här, som hela analysen baserar på. Men samtidigt finns det klart utrymme för en djupare undersökning även om man till exempel tar andra ord i fokus eller andra korpusar för att samla sitt material.

Sammanfattning

Den här bacheloruppsatsen är en kvantitativ analys av förekomst av ord som ursprungligen kommer från arabiska, turkiska, persiska, syrianska/assyrianska i urval av textkorpora.

Huvudsyftet med denna uppsats är att få fram hur ofta och i vilket sammanhang utvalda ord har förekommit i urval av textkorpora. Ämnet var valt av två orsaker: mitt eget intresse för språkändringar och nya företeelser i ett språk samtidigt ökande av antalet invånare i Sverige som räknar sig att ha ett annat modersmål än svenska.

Det analyserande materialet består av 20 utvalda ord som alla har blivit undersökta med hjälp av Språkbankens korpusverktyg Korp. Genom analysen visades det hur många gånger ett ord förekommit i de utvalda korpusexterna, vilka sammanhang de mest har använts i, vilket år ordet först förekom i texterna, hur ofta det användes senare efter det först förekom och om det fanns andra förekommande lemmar. För metoder har det använts både kvantitativ och kvalitativ metod för att genomföra undersökningen.

Analysen visade att bland de utvalda 20 orden förekom 6 av dem i korpusexterna totalt mer än 1000 gånger. Mest förekommande ord var *guzz*, *habibi*, *gäri*, *jalla*, *walla*, *shuno*.

Sammanhanget, där ord mest förekom var Twitter, näst mest populära sammanhanget för ord var två diskussionsforum „Flashback“ och „Familjeliv“. Det mest etablerade ordet i skriftspråket var *guzz*, som förekom även med svenska böjningssuffix – *ar* och *-en* i de utvalda textkorpora. Det enda ordet som ganska sent förekom för första gången i texterna var „gäri“ men efter ordet varit med i Nyordslistan 2014 har dess förekomst ökat kraftigt i texterna. Ordet hade mest träffar år 2017.

Det är en begränsad undersökning som innehåller endast 20 ord från ovannämnda främmande språken. För att mer generella slutsatser kunde dras om hur främmande ord används i svenskan behövs det en större undersökning. Det är möjligt att använda samma metod och ta andra ord i fokus eller undersöka vidare i vilket kontext de har använts.

Referenser

- Auer, P. (ed.) (1998) *Code-Switching in Conversation: Language, Interaction, and Identity*. London: Routledge .
- Fraurud, K. (2003). *Svenskan i Rinkeby och andra flerspråkiga bostadsområden*. Språk i Norden Sprog i Norden, 77
- Kennedy, G. (2014). *An introduction to corpus linguistics*. Routledge.
- Kotsinas, Ulla-Britt (2005). *Invandrarsvenska*, Uppsala: Hallgren & Fallgren
- Kotsinas, U. B. (2007). *Ungdomsspråk*. Uppsala: Hallgren & Fallgren.
- Lindberg, I (2003) *Myter om tvåspråkighet*, Språk I Norden, Novus forlag
- Meyer, C. F. (2002). *English corpus linguistics: An introduction*. Cambridge University Press
- McEnery, Tony, and Andrew Hardie (2012) *Corpus linguistics: Method, theory and practice*. Cambridge University Press,
- Parkvall, M. (2009). *Sveriges språk-vem talar vad och var?* Stockholms Universitet
- Parkvall, M. (2016). *Sveriges språk i siffror: Vilka språk talas och av hur många?*. Institutet för språk och folkminnen, Språkrådet/Morfem
- (Hämtad från <https://modersmalsundervisning.wordpress.com/2016/04/06/sveriges-storsta-sprak-2012/>)
- Rawoens, G. (2003). *Korpuslingvistik och kontrastiv språkbeskrivning. Ett svensk-nederländskt projekt*. Tijdschrift voor Skandinavistiek, 24(1), 1-20
- Språkbankens websida. Hämtad 16.05.2018
- <https://spraakbanken.gu.se/swe/om>
- Statistas websida. Hämtad 16.05.2018
- <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/>
- Stenström, Anna-Brita, Ulla-Britt Kotsinas & Eli-Marie Drange. 2000. *Ungdommers språkmöter*. (Nord 2000: 26.) København: Nordisk Ministerråd

Källor för ord

- Svensk Slang Ordbok websida. Hämtad 5.05.2018
- <http://sv-slang.wikidot.com/invandrarsvenska>
- P5 STHLM. (2016) Tio vanliga arabiska ord. 5 april. Hämtad 5.05.2018
- <https://sverigesradio.se/sida/artikel.aspx?programid=2842&artikel=6404221>
- Slangopedia websida. Hämtad 10.05.2018
- <http://www.slangopedia.se/ordlista/?ord=syrianska>

Loudiyi, M. (2016) 35 ord som bara äkta ortenbarn förstår. GP, 11 juni. Hämtad 10.05.2018

<http://www.gp.se/nyheter/sverige/35-ord-som-bara-%C3%A4kta-ortenbarn-f%C3%B6rst%C3%A5r-1.2586616>

Svensk-dansk ordbok websida. Hämtad 10.05.2018

<http://sprogbro.org/herkomst/7?page=2>

Bilagor

Bilaga 1

Ord	Akade- miska	Finland- svenska	Skönlit- teratur	Sociala medier			Tidnings- text	Webb- nyheter	Wiki- pedia
				Blogg	Forum	Twitter			
guzz	1	0	0	42	883	1599	8	20	3
habibi	5	0	0	22	691	1468	6	3	6
gäri	0	0	0	11	29	1625	0	0	0
jalla	0	0	0	104	634	716	7	3	0
walla	0	0	0	14	330	1016	0	0	0
shuno	0	0	0	2	76	1254	1	0	0
suedi	0	0	0	0	90	884	1	1	0
haffla	0	0	0	3	9	277	1	0	0
benim	0	0	0	0	11	254	0	0	0
shaytan	0	0	0	2	105	79	0	0	0
salaam	0	0	1	12	132	34	2	1	1
tamam	0	0	0	8	7	157	0	0	0
flos	0	0	0	1	88	64	0	0	0
parra	0	0	0	4	57	41	8	0	14
albi	0	0	0	1	42	36	0	1	5
araba	0	0	0	0	7	78	0	0	0
abri	0	0	0	0	34	32	0	0	5
dunya	0	0	0	3	19	41	0	0	0
balesh	0	0	0	0	1	1	0	0	0
brakkam	0	0	0	0	0	2	0	0	0

Bilaga 2

ORD	1997	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
guzz	1	2	0	7	4	20	65	118	70	87	80	84	161	248	306	165	402	369	361
habibi	2	0	1	1	1	4	35	46	141	134	139	21	59	104	141	88	642	488	143
gäri	0	0	0	0	0	0	0	0	0	0	0	0	3	13	59	38	306	538	708
jalla	0	1	4	5	3	11	20	52	73	54	61	98	157	77	118	77	229	280	129
walla	0	0	0	0	1	4	6	13	16	9	20	86	63	61	132	72	356	366	155
shuno	1	1	0	2	0	0	1	1	1	2	3	14	9	16	45	37	459	477	265
suedi	0	0	0	0	0	0	3	0	0	5	9	11	6	12	15	16	324	440	135
haffla	0	0	0	0	0	0	0	0	2	0	6	5	13	6	21	8	80	119	30
benim	0	0	0	0	0	0	1	0	2	0	3	1	1	2	5	4	60	117	69
shaytan	0	0	0	1	0	2	3	13	21	3	12	13	6	19	4	8	21	37	23

ORD	1997	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
salaam	0	0	0	0	0	0	4	12	15	8	17	14	16	20	13	14	6	18	5
tamam	0	0	0	0	0	2	1	0	2	0	0	0	1	7	10	18	57	54	20
flos	0	0	1	1	0	1	2	2	2	4	11	22	7	13	17	9	25	20	13
parra	0	0	0	0	1	0	8	10	7	8	7	2	6	8	9	7	19	12	4
albi	0	0	0	0	0	0	1	0	5	2	7	2	4	5	7	3	17	17	10
araba	0	0	0	0	0	0	2	0	0	1	0	2	1	1	9	3	30	30	6
abri	0	0	0	1	0	0	0	3	0	7	4	7	5	3	5	3	11	14	3
dunya	0	0	0	0	0	0	0	0	1	0	2	1	2	4	3	8	14	18	10
balesh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
brakkam	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Resümee

Tartu Ülikool

Maailma keelte ja kultuuride kolledž

Skandinavistika osakond

Bakalaureusetöö: Guzz, suedi, habibi – kvantitatiivne analüüs sisserändajate keele sõnade esinemisest rootsi kirjakeeles

Autor: Kätlin Kulbin

Juhendaja: Maiu Elken

Kevadsemester 2018

Antud bakalaureusetöö eesmärgiks on uurida sisserändajate keelest pärinevate sõnade esinemissagedust korpuse tekstide valimis. Uuritavateks laensõnadeks on sõnad, mis tulenevad araabia, türgi, süüria ja pärsia keeltest. Uurimustöö on põnev ja relevantne just selle poolest, et sarnaseid uurimusi, kus aluseks on võetud korpuse tekstid, selleks et uurida kui palju ja millises kontekstis mingit sõna esineb just kirjakeeles, ei ole väga laialdaselt tehtud. Võimalik on leida sarnase teemakäsitlusega töid, kus uuritavaks aluseks on võetud kindlad allikad nagu raamatud või blogid, kus uuritakse sõnade erinevaid esinemisvorme nende kindlate tekstide põhjal. Kuid üldist esinemissagedust uuritud ei ole. Vähemalt nii palju kui otsida eelnevaid bakalaureusetöid või siis uuringuid, neid ei leidnud.

Bakalaureusetöö on jagatud nelja suuremasse ossa. Esimeses sissejuhatavas osas põhjendatakse teemavalikut ja seatakse küsimused, millele analüüsi käigus vastust otsitakse.

Küsimused on järgnevad:

Milliseid väljavalitud sõnu esineb korpusetextide hulgas enim?

Millistes tekstides esineb väljavalitud sõnu enim?

Teises ehk teoreetilises osas tutvustatakse varasemaid uurimustöid, mis on läbi viidud sarnasel teemal. Samuti kirjeldatakse teooriaid, mis uurimustöös aluseks on võetud.

Kolmandas osas tutvustatakse metoodikat, mida on kasutatud analüüsi läbiviimiseks samuti esitletakse analüüsitava materjali, ehk väljavalitud sõnu, millel analüüs baseerub.

Analüüsiosas selgitatakse, milliste küsimuste abil iga sõna on analüüsitud materjali kogumiseks. Samuti tuuakse välja iga sõna juures välja vastused antud küsimustele ning

antakse mõningad näited sellest, kuidas see sõna oli tekstides esinenud. Tulemuste ja arutelu osas aga antakse ülevaade analüüsi tulemustest. Tuuakse välja milliseid sõnu esines väljavalitud korpusetekstides kõige sagedamini. Samuti millistes kogumikes esines neid sõnu kõige rohkem. Hiljem arutletakse saadud tulemuste üle ning tuuakse välja mõned selgitused ja võimalikud põhjendused analüüsiosas selgunud tulemustele.

Analüüsi käigus selgus, et väljavalitud sõnadest esinesid *guzz*, *habibi*, *gäri*, *jalla*, *walla*, *shuno* korpusetextides rohkem kui 1000 korda. Kõige levinum sõna oli *guzz*, mis tuleneb türki keelest ja mis esines korpusetextides 2 556 korda. Kõige vähem levinumad sõnad olid *balesh* ja *brakkam*, sõnad tulenevad vastavalt türki ja kurdi keeltest. Neid esines mõlemaid tekstides ainult kahel korral. Kõige rohkem esines sõnu Twitteris, teisena esines sõnu kõige rohkem internetifoorumide tekstide hulgas.

Lihthitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Kätlin Kulbin,

1. annan Tartu Ülikoolile tasuta loa (lihthitsentsi) enda loodud teose

Guzz, suedi, habibi - en kvantitativ analys av ord från invandrarspråk i svensk skriftspråk/
Guzz, suedi, habibi - kvantitatiivne analüüs sisserändajate keele sõnade esinemisest rootsi
kirjakeeles

mille juhendaja on Maiu Elken,

- 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihthitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.



Tartus, **21.05.2018**